An Analysis of the Impact of Configuration Changes to the Learning Curve for Department Of Defense Aircraft Acquisition Programs Substantially Into Production

THESIS

Candice M. Honious, Civilian, USAF

AFIT-ENV-MS-16-M-256

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-16-M-256

An Analysis of the Impact of Configuration Changes to the Learning Curve for Department Of Defense Aircraft Acquisition Programs Substantially Into Production

THESIS

Presented to the Faculty

Department of Aeronautics and Astronautics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Cost Analysis

Candice M. Honious, MS

Civilian, USAF

March 2016

AFIT-ENV-MS-16-M-256

An Analysis of the Impact of Configuration Changes to the Learning Curve for Department Of Defense Aircraft Acquisition Programs Substantially Into Production

Candice M. Honious, MS

Civilian, USAF

Committee Membership:

Dr. John Elshaw
Chair

Lt Col Brandon Lucas
Member

Mr. Anthony Munafo
Member

AFIT-ENV-MS-16-M-256

Abstract

This thesis examines the impact of configuration changes to the learning curve when implemented during production. Prior learning curve research has studied various learning curve models and impacts to the learning curve when other than constant and stable production exists. However, this research is one of the first empirical studies of the impact to the learning curve slope and touch labor hours when production is continuous but a configuration change interrupts the learning process.

This research effort analyzed one joint service and three Air Force aircraft programs. The analysis discovered the learning curve slope after a configuration change is different from the stable learning curve slope pre-configuration change. The differently configured units were found to be statistically different from one another, which may be due to the unstable slope, given that the labor hours per unit are partially a function of the learning rate. The significant difference between the configurations provides statistical support that the new configuration should not be estimated with the learning curve equation of the prior configuration. The research also discovered the post-configuration slope is always steeper than the stable learning slope. Therefore, estimating the new configuration based on the slope of the units pre-configuration change will result in over-estimation, but an initial estimate with a stable slope and no anticipated changes will under-estimate the production hours once a change is required. The steeper slope decreased with each subsequent unit until the slope stabilized. Possible explanations and implications of all analysis and results as well as suggestions for future research are provided.

Acknowledgments

Completing the coursework in the Cost Analysis degree curriculum and this thesis effort has pushed me in ways I never could have imagined. I sincerely appreciate the efforts of my entire thesis committee, without whom I could not have completed this academic program or thesis project. The guidance, insight, and challenging discussion from my thesis committee member and sponsor, Mr. Anthony Munafo, immeasurably improved my thesis project and professional aptitude. I also appreciate the guidance, feedback, and extreme patience of my thesis advisor, Dr. John Elshaw, and my program advisor, Lieutenant Colonel Brandon Lucas, whose time and discussion were instrumental in my research. I am personally and professionally grateful to all of you.

Most importantly, my deepest gratitude and love is for my wonderful husband. Our time together was sacrificed so I could dedicate myself to this program and especially my thesis project. Your unwavering support and reassuring words kept me focused and encouraged. Thank you for supporting me, believing in me, and being there for me. I could not have done it without you. I love you.

Candice Honious

Table of Contents

List of Figures

List of Tables

An Analysis of the Impact of Configuration Changes to the Learning Curve for Department Of Defense Aircraft Acquisition Programs Substantially Into Production

I. Introduction

Planes became obsolescent as they were being built. It sometimes took five years to evolve a new combat airplane, and meanwhile a vacuum could not be afforded. . . . I also had trouble convincing people of the time it took to get the "bugs" out of all the airplanes. Between the time they were designed and the time they could be flown away from the factory stretched several years … You can't build an Air Force overnight.

- General Henry H. ("Hap") Arnold, 1949

General Issue

Aircraft manufacturing has existed for nearly a century, yet many of General Arnold's observations regarding aircraft procurement issues during WWI still exist today: a long acquisition process, a constrained fiscal environment, and configuration changes introduced during production. The Budget Control Act of 2011 subjected the Department of Defense (DoD) to a more fiscally constrained and financially conscious environment than ever before, juxtaposed with a demand for new aircraft programs of almost every type. As an increasing number of programs are terminated, with budget overruns as a contributing factor, managers at every level in the DoD are expected to ensure the Department's shrinking budget is being used in the most effective way. The increased scrutiny adds greater emphasis to the accuracy of program office cost estimates given that an approved program cost estimate supports every major aircraft acquisition program funded by the Department.

A vital input to the cost estimate for a production program is the assumed learning curve slope for the program. The learning rate is also a major factor in production contract negotiations and has a direct impact on the procurement costs and the contract amount (Hall, 2001:8). The learning curve often depicts the learning phenomenon that occurs in manufacturing. Learning is defined as a constant percentage reduction of the required touch labor hours (or costs) to produce an individual unit as the quantity of units produced doubles (Yelle, 1979:302)—as the number of units produced doubles, the number of hours required to produce a single unit decreases by the learning curve rate. Learning is both the conceptual and the physical learning of a physical process (Watkins, 2001:18). The learning curve for a program is generally considered stable once the program is substantially into production because the manufacturer and laborers have produced enough units to learn the most efficient production process. However, intuitively and through past research, it is known that learning is disrupted by changes in production and only the production of additional units can recover the lost learning (Watkins, 2001:18). It is critical to capture the change in the learning rate due to production modifications to better estimate DoD program costs.

The idea of learning in a production environment is well established. T.P. Wright first published the learning curve phenomenon in early 1936. Wright observed that in a manufacturing environment, as the cumulative quantity of units produced doubled, the cumulative average cost decreased at a constant rate (Wright, 1936:124-125; Yelle, 1979:302). During World War II, government contractors investigated the usefulness of the learning curve concept to predict labor hours and cost requirements for aircraft and ship construction projects. The private sector went on to adopt the learning curve theory into practice shortly thereafter. Although learning curve theory has evolved and has been referred to by different names in the decades following Wright's report, including the experience curve, the progress curve, and the improvement curve, Wright's model remains one of the models most widely used by manufacturers to predict labor hours and costs (Yelle, 1979:303-304; Badiru, 192:176.).

Although Wright's findings postulated a constant learning environment, researchers have not ignored the idea that constant learning may not exist on a continual basis in a manufacturing environment. In fact, the ideas of regressed and lost learning have been widely studied. Research studies support that a break in production creates an environment of relearning because the labor resources have stopped working, at least on the same project, and will be less efficient at manufacturing when production restarts (Anderlohr, 1969:16-17). George Anderlohr (1969) developed a method to determine the cost of lost learning due to production breaks; an overview of this method will be discussed in Chapter II.

In addition to production breaks, instances also exist when a major configuration change occurs during production and disrupts the learning process. In this situation, the new configuration is immediately incorporated into the next units on the production line; the units already produced are retrofitted a later time. Intuitively, the units with the configuration change should initially have a different learning rate than the units without a change because the manufacturers must learn how to incorporate the change into the production process. However, because the learning rate for the new configuration is unknown, DoD program offices generally do not treat the reconfigured units with a different learning rate. As a result, the program often experiences substantially different hours/costs for the newly configured production units than the learning curve projected. A configuration change in a production program does necessitate learning for the contractor, and the impact to learning attributable to the configuration change should be understood by all levels of the DoD acquisition community. Wright (1936) understood this limitation to the learning curve theory application even in the infancy of the idea:

> The tremendous cost of changes introduced into a production order during construction is too well known to require emphasis. This cost is involved, not only in shop delays, but in the engineering expense of re-designing. It is appreciated that in a rapidly moving art such as aviation, changes are more or less inherent… In using the curve developed in this paper, it should be recognized that the factors derived are based on the assumption that *no major changes will be introduced during construction*. (Wright, 1936:124)

Problem Statement and Research Hypothesis

Current DoD program office cost estimating assumes a stable rate of learning once a program is substantially into production. However, intuitively, a configuration change introduced into the production line will initially disrupt the learning effect. This study will research two main questions to address the implications when a configuration change occurs during production:

1. Is there an impact to the learning curve slope when a configuration change is introduced to the production line? Specifically:
    a) What is the learning curve slope for each new configuration;
    b) Are the production segments for each configuration significantly different; and
    c) What is the difference between the hours predicted based on the prior configuration and actual hours for each segment?
2. How many units of the newly configured aircraft are produced before the contractor recovers the stable learning rate?

The first research question leads to a single testable hypothesis:

Hypothesis 1: Is the mean amount of labor hours prior to a configuration change the same as the mean amount of labor hours subsequent to a production change?

H_0: Mean labor hours $_{\text{prior to configuration change}}$ = Mean labor hours $_{\text{post configuration change}}$
H_a: Mean labor hours $_{\text{prior to configuration change}}$ $\neq$ Mean labor hours $_{\text{post configuration change}}$

If the analysis results fail to reject the null hypothesis, this would indicate the data points come from the same population and a configuration change did not have a significant impact to the learning during production. If the analysis rejects the null hypothesis, this would indicate the opposite, the data points representing different configurations come from different populations and that a configuration change did have a significant impact to the learning during production. If the results support rejecting the null hypothesis, using the prior learning curve equation is inappropriate to predict the hours of the new configuration because the units come from different populations. The second research question does not require a hypothesis test.

Methodology

Data will be collected from aircraft program offices at Wright-Patterson Air Force Base and analyzed to determine the change in touch labor hours or costs incurred by the contractor when a configuration change was introduced to the production line. Regression techniques will evaluate and compare the learning curve slopes for the aircraft units both prior and subsequent to the configuration change(s). The regression analysis will also explain if the actual labor hours between the configurations are statistically different, implying an adjustment to the learning curve slope is necessary to account for the changes. Chapter III provides an in-depth approach to the methodology.

Assumptions/Limitations

This research has a broad scope and is only limited by the availability of data at the necessary level of detail. This study is analyzing configuration changes, data that can only be obtained from aircraft program offices versus pulling data from a DoD database. To analyze the impact to the touch labor hours and learning curve slope after a configuration change occurs, the data are needed by individual aircraft unit—not just for the lot in its entirety. To determine the impact based on type of configuration change, the data are needed by work type classification—not just for the unit at a top level. Chapter III of this report contains a further description of the data requirement subtleties.

Implications

If the data analysis produces significant results, the final aspect of the study will be to determine if developing a Cost Estimating Relationship (CER) is feasible. This CER would project the impact to a cost estimate for unplanned configuration changes in a specific program. Such a factor/CER could empower the DoD during contract negotiations for aircraft lot buys. The contractor and program office are both aware that configuration changes will most likely occur between awarding the initial contract and the delivery of the final aircraft. Based on this empirical data, the contractor adds dollars to the production estimate to cover the cost of configuration changes, and the DoD is currently unable to assess the reasonableness of such estimates. If a method can be developed, the results will improve the DoD's ability to negotiate aircraft production contracts.

Review

This chapter provided the rationale for analyzing the topic and outlined the research objectives. Chapter II describes past and present research efforts regarding the concept of learning. Chapter III details the methodology used to analyze the aircraft data. Chapter IV summarizes the results of the data analysis. Chapter V provides the study conclusions and recommendations for further research.

II. Literature Review

Introduction

The idea of increasing efficiency while repeating a task, especially in a high touch-labor environment like manufacturing, has been around for decades. In 1936, T.P. Wright published the first learning curve model, which evaluated the cost of aircraft, but manufacturers in both the public and private sectors realized this learning phenomenon applied to most production environments. Since Wright's discovery, several learning curve models have evolved from Wright's original concept, but the premise behind each model remained the same: a production labor resource will take less time to complete a task the more often the laborer repeats that task without a break in work performance. When a disruption occurs that impedes the laborer's ability to repeat the same task, the laborer's efficiency and the learning rate are both impacted.

The purpose of this chapter is to provide an in-depth review of current learning curve theories and methodologies, especially those most often used within the DoD. This chapter will also examine past research in the areas of lost learning in a production break environment, the more contemporary learning curve research of forgetting rather than learning, and the less explored area of split learning curves.

Relevant Research

First, the relevant research begins with the genesis of the learning curve concept. Next, the literature review continues with the evolution of learning methodologies. Finally, the research overview ends with the related concepts of lost learning due to production breaks, forgetting and its causes, and accounting for the addition of new work during manufacturing.

Learning Curve Theory Conception

T.P. Wright (1936) identified the learning phenomenon in a manufacturing environment and published the first learning curve model in the 1930s. Wright identified that as production laborers repeated the same task in aircraft manufacturing, the laborers learned from their prior repetitions and became increasingly more efficient at the task (Wright, 1936:124). Manufacturers in all industries quickly adopted the concept.

After graphing the variation of labor cost against aircraft production quantity, Wright identified that the learning curve was of an exponential form, shown in Equation 1 below. This model, often referred to as Wright's Learning Model, shows mathematically that as the cumulative production quantity doubles, the cumulative average production cost decreases at a constant rate.

$$y = ax^b \tag{1}$$

Where

y = the cumulative time or cost after producing x number of units
a = hours required to produce the theoretical first unit
x= cumulative unit number
b = log R/log 2 (the learning index)
R = learning rate

Wright used this formula to develop the 80% curve, based on the aircraft data he plotted. Wright believed aircraft manufacturers would observe the 80% learning rate as the cumulative quantity produced doubled (Wright, 1936: 124-125). While civilian aircraft assembly still has an expected learning curve slope of around 80%, learning curve slope values can vary across and even within industries (ICEAA Module 7, 2013:59).

Figure 1 depicts Wright's 80% learning curve, based on a first unit cost of $100 thousand. The figure shows that as the cumulative number of units produced doubles, the cumulative average cost decreases by 20%.

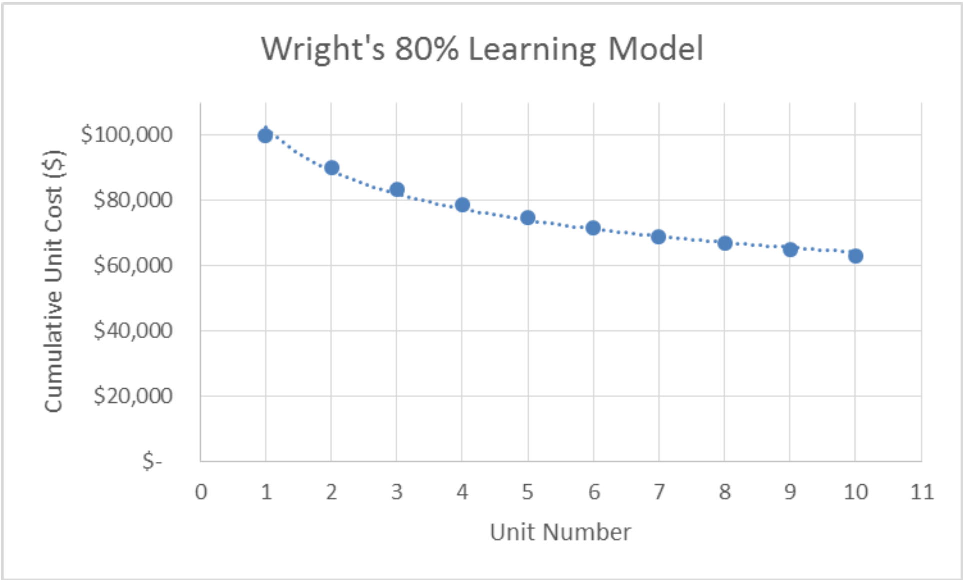

Figure 1: Example of Wright's 80% Learning Curve

Wright also published the log-linear form of his equation, show in Equation 2, which transforms the plotted data from a curved line into a straight line (still referred to as a learning curve) (Wright, 1936:124). Practitioners often transform production data into this form, because statistical regression can then be used to plot a line that best fits among all of the data points when in this form.

$$\ln y = \ln a + b \ln x \tag{2}$$

Figure 2 depicts the transformation of Wright's 80% model into log-linear form. When a constant learning rate exists, the log-linear learning curve is a straight line. When the log-linear curve data approximate a straight line, laborers are achieving constant efficiency at the learning rate.

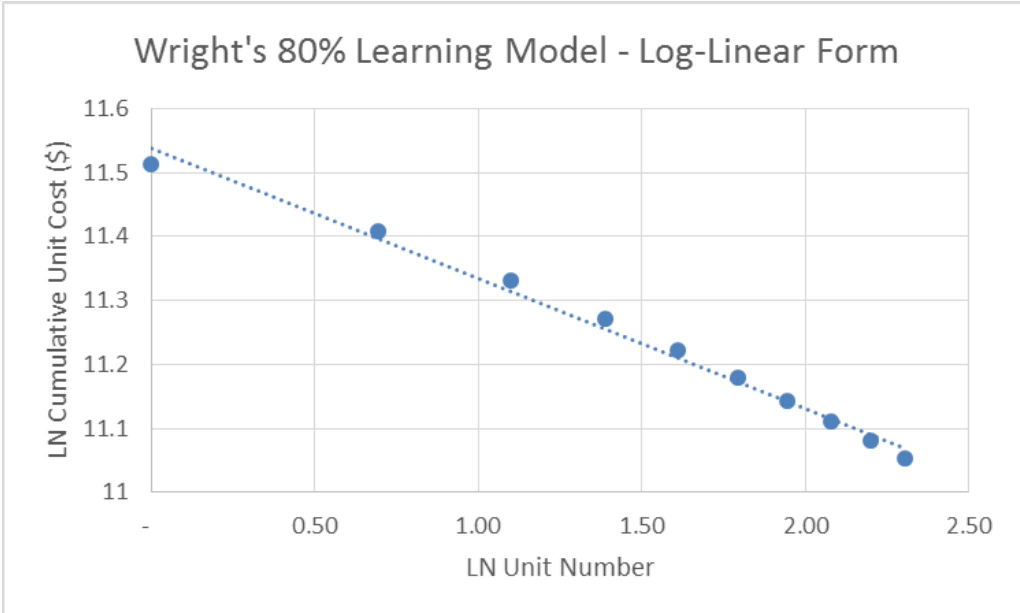

Figure 2: Log-Linear Learning Curve Example

J.R. Crawford repurposed Wright's model during his time at Lockheed Martin, using the same equation and underlying theory, but defining the X and Y variables differently. Crawford defined the X variable as the individual unit number and the Y variable as the individual unit cost instead of the cumulative value for both variables (ICEAA Module 7, 2013:31). Identifying data by individual units rather than cumulative production enables easier detection of production units manufactured more or less efficient than others are, which can then be studied as to why. Figure 3 shows an example of Crawford's model with an 80% learning curve and first unit cost of $100 thousand.

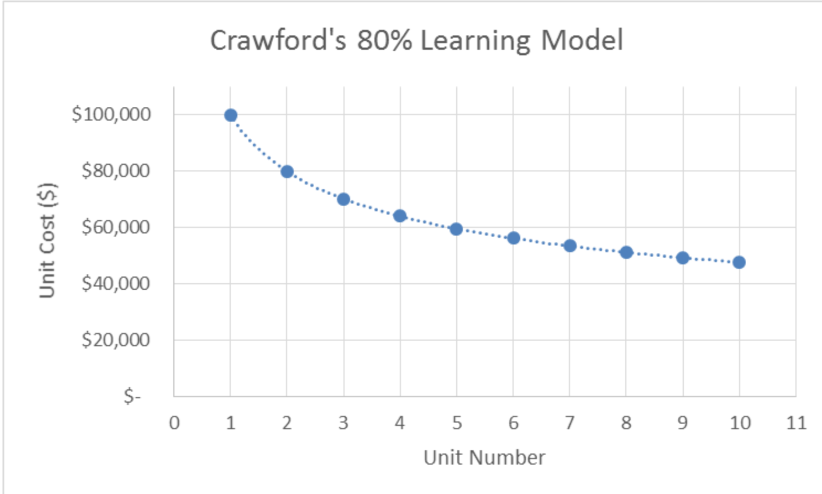

Figure 3: Example of Crawford's Learning Curve

The same power form is evident in both Wright and Crawford's models, but as Figure 4 shows, predicting the unit cost (Crawford) results in a lower per unit cost prediction than the cumulative average unit cost (Wright) theory. This is intuitive because Wright's model presents data as average, which smooths the impact of any specific data point. While the advantages of observing the data in unit space are obvious, the cumulative average cost method is still more popular than the unit cost method (Badiru, 1992:176).

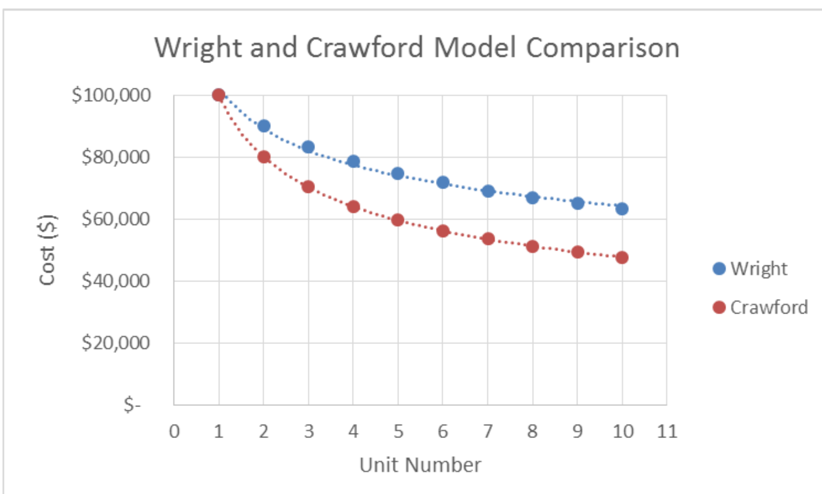

Figure 4: Wright and Crawford 80% Curve Comparison

The learning curve theory is especially important to the aircraft industry because aircraft manufacturing is a high touch labor environment. Henneberger and Kronemer (1993) identified three factors that contribute to the low-automation environment in this industry: (1) the aircraft industry produces a customized product, (2) the unit volume produced is extremely low compared to most manufacturing activities, and (3) the aircraft industry produces a complex product (26). These three factors do not incentivize the aircraft industry to invest in laborsaving machinery, because a highly skilled workforce is more cost effective in an environment that demonstrates these traits; DoD aircraft manufacturing is no exception, and is probably more prone to these three factors than commercial aircraft manufacturing.

Learning Curve Theory Evolution

Wright's learning curve theory drastically influenced the manufacturing industry after World War II. Numerous learning curve models have since been developed to more accurately represent what manufacturers uniquely observed taking place during production. Figure 5 graphs the most well known models and illustrates the differences depending on what cumulative production unit number is being observed. The five most well known models are: the log-linear model (Wright/Crawford), the plateau model, the Stanford-B model, the DeJong model, and the S-model (Yelle, 1979:304). The following sections describe the distinctive qualities of each model type.

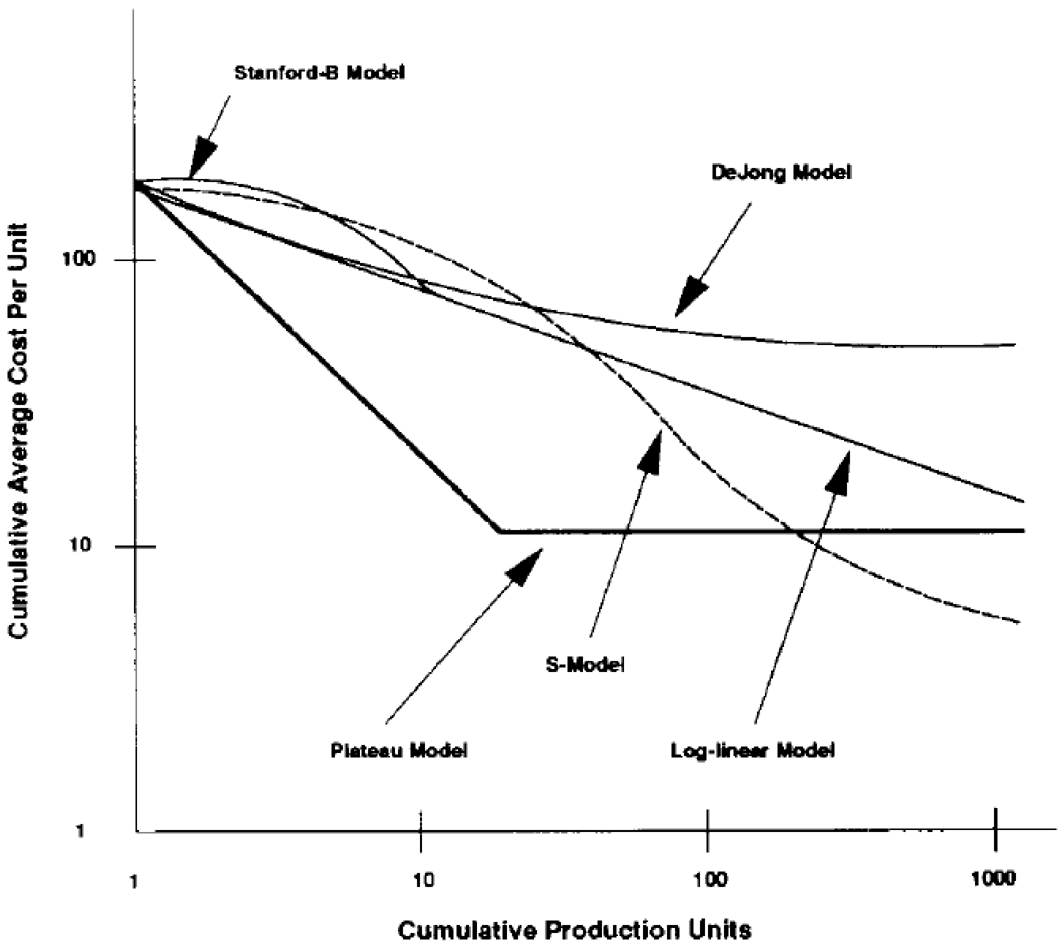

Figure 5: Comparison of Learning Curve Models on a Log Scale (Badiru, 1992)

Plateau Model

Levy (1965) suggested Levy's Adaptation Function, also known as the plateau model, and the main distinction of this model is the use of a constant to flatten (plateau) the learning curve as the number of units produced becomes large (Badiru, 1992:179). The idea behind this constant is that learning does not continue indefinitely, and as the number of units produced reaches a certain level, any change in the learning curve (decrease or increase) is ignored and the curve flattens.

Stanford-B Model

The Stanford Research Institute conducted a study (1956) on behalf of the DoD that led to the development of this model. An important distinction in this model is adding a constant value between one and ten to adjust the production unit (X variable) value being estimated. The constant value is equal to the number of units of prior experience before the first unit acceptance occurs. If there are no prior equivalent units, the Stanford-B model reduces to the conventional Crawford model because the constant's value in the equation is zero. Prior aircraft production research has found the addition of this constant to account for the most relevant prior learning has reduced statistical error in a regression model (Badiru, 1992:178).

DeJong Model

DeJong (1957) distinguished from the conventional log-linear model with the addition of an incompressibility factor between zero and one to account for the proportion of task activities between manual and machine operations. An incompressibility factor of zero indicates entirely manual operations while an incompressibility factor of one indicates entirely machine. The one also indicates that there is zero cost improvement (learning) possible (Badiru, 1992:179).

S-Model

Carr (1946) is credited with developing the S-Curve model. Figure 5 depicts the model in log space, but in unit space, the model takes the form of the letter "S," making this model easy to recognize. An important distinction in this model is that a gradual start up exists as production begins. The production process and its laborers becoming more efficient over time as more units are produced cause this transitory ramp-up state (Badiru, 1992:178).

While this research study will not focus on applying the various learning curve models, it is important to point out that different models have been developed over time and are used in practice based on which model most accurately fits a manufacturer's learning environment. The five models just discussed are the five most popular models, but the list of learning models aforementioned is not exhaustive.

Production Break and Lost Learning

As the learning curve theory has evolved, researchers and practitioners have investigated the impact to the learning rate, when other than constant production exists. George Anderlohr (1969) is credited with developing a model to determine the additional hours/costs that result from a break in production. Anderlohr (1969) defined a production break as, "the time lapse between completion of a contract for the manufacture of certain units of equipment and the commencement of a follow-on order for identical units" (16). A break in production results in increased hours and costs, because the laborers are no longer performing their tasks on a constant repetitive basis, and the laborers become less efficient (have a loss of learning) during the production break timeframe (Anderlohr, 1969:16-17).

Anderlohr's method identified five factors that contribute to a loss of learning: production personnel learning, supervisory learning, continuity of production, improvement of special tooling, and improvement of methods. Each factor was weighted based on the specific company and industry. The amount of personnel and manufacturing process that remains (or is lost) after the production break is estimated to develop a total percentage of learning lost. The total learning lost during the production break is used to regress back up the learning curve before the production of the next lot begins. Regressing back up the learning curve to account for the break will more accurately predict the total hours for the next lot given that re-learning must take place before the laborers will become as efficient as before the break (Anderlohr, 1969:16-17).

Studies have also discovered that lost learning can be a result of forgetting at times other than during a production break (which is considered scheduled forgetting). Two other instances when forgetting can occur are: 1) at random due to the inability to continue work (e.g. machine breakdowns), and 2) based on a natural process (e.g. aging workforce) (Badiru, 1995:780). Badiru goes on to conclude that, "whenever interruption occurs in the learning process, it results in some forgetting." The amount of forgetting is a function of both the length of disruption and the initial performance level (Badiru, 1995:780).

Given that the initial performance level will depend on the individual laborer, the amount forgotten will also depend on the individual laborer. A laboratory study conducted to study the effects of forgetting discovered three important findings regarding learning and forgetting. First, laborers do not have insight into their own memory. Second, individual learning is highly correlated to the amount of time taken to complete the first production unit. Finally, learning, forgetting, and relearning are not necessarily the same rate, but rather a function of individual skill levels and the initial learning (Bailey, 1989:340).

Lam et al. agree that some forgetting will occur whenever a production interruption occurs but postulate that forgetting initially occurs as a rapid decrease in performance that gradually plateaus (Lam et al., 2001:412). Figure 6 illustrates this forgetting phenomenon.

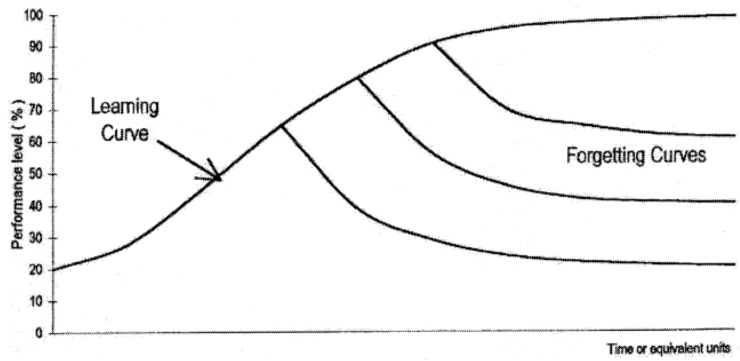

Figure 6: S-shaped Forgetting Curve (Lam et al., 2001)

The rapid decrease and amount of time before plateauing depends on the number of successive units completed without disruption prior to the production interruption, as other research has suggested. Total forgetting occurs only after sufficiently long (undefined) breaks in production. As long as total forgetting has not occurred, Figure 7 depicts the idea of learning, some forgetting, and relearning (Lam et al., 2001:414).

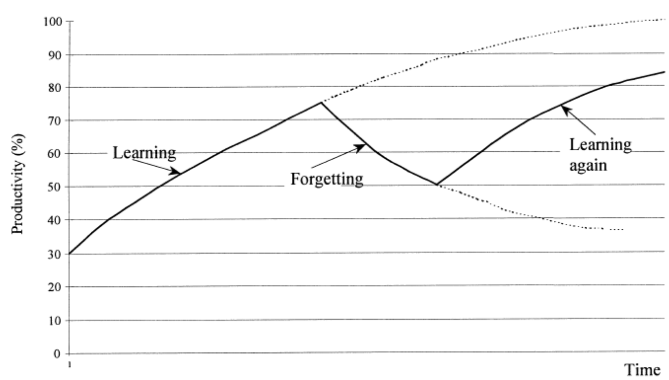

Figure 7: Learning-Forgetting-Learning Curve Example (Lam et al., 2001)

While this study will not focus on incorporating the ideas of lost learning and forgetting into any calculations, forgetting and relearning and the overall principle that learning does not continue at a constant rate for the life of production is a growing area of importance in the learning architecture.

Additional Work Theory

A similar circumstance to the production break theory, that has a similar result, is the idea of new learning, when manufacturing is interrupted with a major configuration change to the production unit. When the unit being manufactured is changed, the laborers must adjust their processes to learn how to correctly produce the newly configured unit. Historically, adjusting the learning curve to account for the impact due to configuration changes is referred to splicing or splitting the curve, although little research has been done in this area with empirical data. The theory of splitting the curve provides rationale to split the curve between units of different configurations (pre- and post-configuration change) because the latest production unit usually provides the greatest estimate for future production units (Dahlhaus and Roj, 1967:16).

One possible method to adjust the learning curve to account for the additional learning has been documented by the International Cost Estimating and Analysis Association (ICEAA), formally known as the Society of Cost Estimating and Analysis (SCEA). This method adds the new learning curve (post-configuration change to the old learning curve, starting with the first unit after the configuration change was introduced (ICEAA Module 7, 2013:77), as shown in Equation 3.

$$Y = a_1 X^{b_1} + a_2 (X - L)^{b_2} \tag{3}$$

Where

Y = the unit cost
a_1 = original T1 value
a_2 = new work T1 value
x = current unit number
L = last unit before addition of new work
b_1 = exponent for original learning curve slope
b_2 = exponent for new work learning curve slope

Evolving this idea to identify the learning rate and cost impacts due to configuration changes is a basis for this research.

While numerous learning curve methods and theories have evolved from Wright's discovery, the Air Force directs its cost analysts to use the unit curve and cumulative average formulations, both of which follow Wright's original exponential learning curve model (AFCAH, 2007:4). Both of these formulas ignore the important reasons to vary the traditional learning curve discussed in this section—different manufacturers sometimes observe learning that is demonstrated better by a model other than the log-linear model, and adjustments to the learning rate may be necessary if other than constant production occurs. This research will try to build upon existing models to identify the proper adjustment to the learning rate when a configuration change cuts into the production line.

Summary

Chapter II creates the foundation for the research by providing a basic historical overview of learning curve concepts for the past 80 years and identifying an apparent gap that this research will attempt to address. Chapter III describes the methodology used in this research study to understand the impact of configuration changes to the learning curve of DoD aircraft acquisition programs.

III. Methodology

Introduction

The premise of this research is that while the DoD directs the use of Wright's learning curve model (or Crawford's application in unit theory) for cost estimating (AFCAH, 2007:8) as described in Chapter II, there are modern complications in production that need to be addressed to more accurately predict aircraft production hours/costs. While Wright's model arguably remains the most widely implemented learning curve theory today, Wright acknowledged the limitations of his model in his initial publication, and stated, "it should be recognized that the factors derived are based on the assumption that no major changes will be introduced during construction" (Wright, 1936:124).

However, current cost estimators do not adjust Wright's learning curve to account for the major configuration changes that come into the production line because the slope of the newly configured units is unknown. Instead, as changes are introduced every unit is treated the same, i.e. keeping all units running down the same, constant slope. The cost estimators account for the configuration changes in the learning curve analysis only after the contractor provides the actual hours/costs to the program office, which can be a considerable amount of time after the change came into production. Developing a method to adjust the learning curve slope or estimate based on the impact of major configuration changes may lead to increased accuracy of cost estimates and greater ability in contract negotiations, which are both vital concerns during this constrained fiscal environment.

Data Collection

The availability of data that meets the intent of this study determines the effectiveness of this research. Prime contractor aircraft production data was collected from the aircraft program offices and the Air Force Life Cycle Management Center Cost Staff (AFLCMC/FZC) at Wright-Patterson Air Force base. The production data was provided in touch labor hours and production costs and was detailed by production unit or production lot depending on the specific program. Current DoD cost databases do not supply adequate data because the data required for this study is so specific, as will be described later.

Research Design

This study will use the traditional learning curve analysis steps as a guide to begin the research analysis (ICEAA Module 7, 2013:32):

1. Collect and normalize data
2. Scatter-plot data and fit a power trendline
3. Transform into log space
4. Plot data in log space
5. Determine linear equation using regression
6. Determine answers through applied learning curve equation

Based on the research questions this study hopes to answer, steps two and four provide significant information that would be more difficult to discern otherwise.

Creating a scatter-plot of the data provides initial insight into the trends and relationships that may exist in the data set (ICEAA Module 1, 2013:61). A scatter-plot analysis can also emphasize breaks between the data points and that a single power curve for the dataset may be inappropriate. This phenomenon is referred to as a “pseudo learning curve”; a single power curve that is statistically significant but poorly fits the data (ICEAA Module 7, 2013:13). Figure 8 depicts a pseudo learning curve example. The plotted data clearly shows that a single power curve may not represent the data as well as three separate power curve segments; units one through three, units four through seven, and units eight through ten. It is visually evident that the single power curve represents a much flatter learning slope than the steeper learning that is apparent in the three segments when analyzed separately.

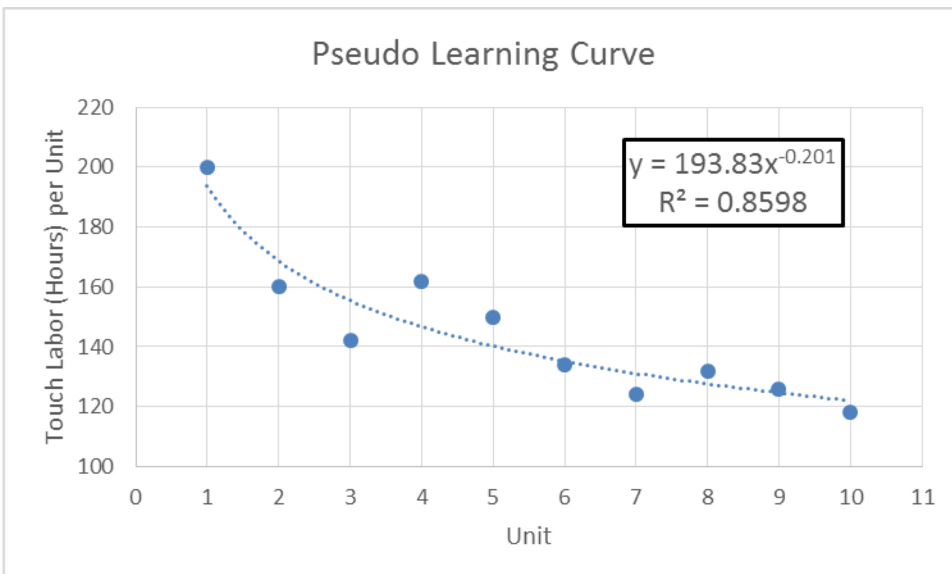

Figure 8: Pseudo Learning Curve Example

A focus of this research study is breaking a pseudo learning curve into the more appropriate multiple power curves to analyze a single data set properly.

Plotting the data in log space visually confirms that the analysis method is appropriate. When plotted in log space, the data should approximate a line. If the data do not approximate a line, the analysis method for that data set may not be suitable (ICEAA Module 7, 2013:22). It is not the intent of this research study to analyze alternative analysis methods, so this visual analysis will be used as a discriminator of which data segments to include in the analysis.

Research Population

The previous section described where all of the data for this study was obtained. The initial data collection resulted in a portfolio of 50+ DoD aircraft production programs of varying platform types. The data required for this study is so specific that the data for only four programs was available to analyze at the time this study was conducted. The three limiting conditions the data had to satisfy to be included in this study were: 1.) At least one identified configuration change must come into the production line during production, 2.) All units must be produced on the same production line, and 3.) The program must be "substantially" into production.

Elaborating on the limiting conditions, this research seeks to isolate the impact due to a configuration change and to reduce the possibility of identifying an impact to the learning curve that is actually due to another factor. Therefore, at least one configuration change must be introduced into the production line. The program offices identified the configuration changes occurring in the data, which current DoD cost databases do not categorize or identify.

Additionally, all units need to be produced on the same production line to allow analysis of a single production process. Learning cannot be analyzed otherwise because different labor resources are used on separate lines and different production processes may be used in separate locations, among other limitations. Finally, the program must be substantially into production for this analysis because it is common for aircraft units late in the development and early in the production processes to be changed and for the production hours to fluctuate due to less efficient processes (i.e. pre-learning). Early production configuration changes are normal in the DoD aircraft acquisition process and unstable learning is expected. Until a stable learning rate is achieved, this analysis cannot isolate the impact due to a configuration change. For the purposes of this analysis, substantially into production is defined as those units considered by the program office to be representative of stable production and exclude any units identified as development or pre-production.

After excluding any programs that did not meet the research conditions, only four data sets remained in the analysis. All data provided in lot groupings were excluded because no detail was available to confirm if a configuration change occurred regardless of what the scatter-plot visually suggested.

The unit theory data included in this analysis includes one joint service and three Air Force aircraft programs. All the data used in the study are provided in hours so the data do not need to be standardized. Due to the proprietary nature of the production data, the program names are not disclosed and will be identified as Programs A, B, C, and D. Only a subset of the joint service program (Program D) labor hour data can be analyzed because only one portion of the aircraft production is completed on a single production line.

The scope of this research project initially included all aircraft classes. After scrubbing the available data, only three classes of aircraft are represented in this study:

Unmanned Air Vehicle, Cargo, and Fighter aircraft. Table 1 summarizes the sample sizes for each program included in this study:

Table 1: Research Population Program Sample Sizes

Program	**Sample Size**
Program A	84
Program B	27
Program C	176
Program D	115

Research Questions and Hypothesis

This research will study two main research questions to address the implications when a configuration change occurs during production:

1. Is there an impact to the learning curve slope when a configuration change is introduced to the production line? Specifically:
    a) What is the learning curve slope for each new configuration;
    b) Are the production segments for each configuration significantly different; and
    c) What is the difference between the hours predicted based on the prior configuration and actual hours for each segment?
2. How many units of the newly configured aircraft are produced before the contractor regains the stable learning rate?

The first research question leads to a single testable hypothesis:

Hypothesis 1: Is the mean amount of labor hours prior to a configuration change the same as the mean amount of labor hours subsequent to a production change?

H_0: Mean labor hours $_{\text{prior to configuration change}}$ = Mean labor hours $_{\text{post configuration change}}$
H_a: Mean labor hours $_{\text{prior to configuration change}} \neq$ Mean labor hours $_{\text{post configuration change}}$

If the analysis results fail to reject the null hypothesis, this would indicate the data points come from the same population and a configuration change did not have a significant impact to the learning during production. If the analysis rejects the null hypothesis, this would indicate the opposite, the data points representing different configurations come from different populations and that a configuration change did have a significant impact to the learning during production. If the results support rejecting the null hypothesis, using the prior learning curve equation is inappropriate to predict the hours of the new configuration because the units come from different populations. The second research question does not require a hypothesis test.

Variables and Statistical Tests

For this research study, the total touch labor hours required for a single aircraft unit's production is the only dependent variable. The learning curve slope contributes to the predicted and actual required production hours. The independent variable is a configuration change occurring during the aircraft production.

The hypothesis involves comparing the mean values pre- and post-configuration change to ascertain if the means are statistically similar. Prior to testing the means, assumptions about the data must be tested to determine if parametric or nonparametric testing is more appropriate. Parametric tests make inferences about the underlying population parameters, whereas nonparametric tests do not depend on the underlying distribution of the population (McClave, Benson, and Sincich, 2014:15-3). Prior learning curve research has found parametric estimating predicts with the smallest range and least dispersion, but has the most bias in the estimated parameters. Nonparametric estimating has been found as the opposite, the least biased estimated slope and first unit cost, but a greater predicted range with more dispersion (Avinger, 1987:41).

A popular parametric analysis test used to compare mean values is the Student's t-statistic. If the results of this t-test are significant, the results indicate that the true mean between the compared values is different (McClave, Benson, and Sincich, 2014:380). The sample populations must meet specific assumptions to use the t-test: the samples must be randomly selected from the population; the samples must be selected independently of each other; and the data must be normally distributed and have equal variances (McClave, Benson, and Sincich, 2014:423-424).

If the underlying population assumption of normality cannot be met, the comparable nonparametric tests are the Kruskal-Wallis Test (comparing three or more populations), the Wilcoxon Rank Sum Test, and the Mann-Whitney U Test (both comparing two populations). These three nonparametric tests compare the medians of the population samples instead of the mean values because a normal distribution is not assumed. The sample data is pooled and ranked as if it came from the same population and if the underlying populations are the same, the ranks should be randomly mixed between the samples. If the underlying populations are different, one sample will have more of the larger ranked values. Results that support rejecting the null hypothesis indicate that the true median between the compared populations is different (McClave, Benson, and Sincich, 2014:15-9).

Nonparametric tests make no assumptions about the distribution form of the data. The only conditions required for the aforementioned nonparametric tests are that the samples are independent and randomly selected and that the distributions are continuous. The only additional requirement for the Kruskal-Wallis test is that each sample includes at least five data points (McClave, Benson, and Sincich, 2014:15-28).

The Wilcoxon test allows a test statistic z-value to be computed and compared to a z-score as with a normal t-test if both sample sizes are greater than ten (McClave, Benson, and Sincich, 2014:15-12). The Wilcoxon test was conducted when both sample sizes were "larger" (greater than ten) and at least one sample size was greater than 30. The Mann-Whitney test was conducted otherwise (i.e., for "smaller" sample sizes, neither greater than 30). The Mann-Whitney and Wilcoxon tests are considered statistically equivalent (McClave, Benson, and Sincich, 2014:15-11), so either test will yield the same results.

This research study will evaluate statistical tests at an alpha value (α) of 0.05. This significance level indicates results are presented with 95% confidence and only a 5% chance exists that the null hypothesis is rejected in favor of the alternative hypothesis when the null hypothesis is in fact true (McClave, Benson, and Sincich, 2014:361). In this analysis, rejecting the null hypothesis provides 95% confidence that the means between the populations are different. If nonparametric analysis is more appropriate for the study, a z-value of 1.96 will be used because it is the corresponding z-value that provides 95% confidence.

Random Sample

A random sample is defined as “a sample selected from the population in such a way that every different sample of size *n* has an equal chance of selection (McClave, Benson, and Sincich, 2014:15). The samples in this study are random because every DoD production aircraft program had an equal chance of being selected.

Independence

Past research of DoD programs has concluded that independence between DoD programs exists if legislation and regulation would affect cost performance similarly for each program and that a multitude of personnel manage the programs and contracts (Searle, 1997:58-59). Given that all DoD aircraft programs fall under the same regulations and legislation, that support personnel with varying experience manage every program, and training, the assumption of independence of the populations is met.

Normal Distribution

The objective assessment of normality will be determined through the Shapiro-Wilk test. The null hypothesis of this normality test is that the data comes from a normal distribution (Searle, 1997:66). If the test p-value is less than 0.05, the null hypothesis is rejected in favor of the alternative, that the population data is not normally distributed. As summarized by Tracht, past research into learning has assumed normality because a normal distribution was frequently observed in industry and because of the feasible range man-hours can assume (Tracht, 1988: 23). However, an objective determination is used in this study to statistically support any findings.

Equal Variance

The assumption of equal variance will be tested by calculating the variance value for each sample and dividing the largest variance by the smallest variance. A rule of thumb states if the resulting value is less than three, the assumption of constant variance is probably met (Ford, n.d.).

Analysis Methods

After applying the statistical methods described in the prior section and determining whether parametric or nonparametric testing is more appropriate, the analysis will progress to the research hypothesis testing. The data will be split into separate segments at each identified configuration change (which should also be evident by an increase in the labor hours). The appropriate statistical test(s) will identify if the segments are statistically similar based on the mean or median labor hour values.

Using the touch labor hours, the learning rate before an identified configuration change and the learning rate after the change will be calculated to address the remaining areas of the first research question. Both calculations will use Crawford's equation $y = ax^b$, which Chapter II detailed; because the data is available in units, a unit analysis is appropriate. In addition, to avoid the smoothing effect and the obfuscation of unit variation a cumulative unit curve can create, the unit learning curve method will provide the most explanatory results of the two methods for the intent of this study (ICEAA Module 7, 2013:14).

The slope will be calculated each time an identified configuration change occurs and not at other instances, even if a pattern change is evident in the scatter-plot of the data. The learning curve equation of a segment will forecast the touch labor requirements of the successive production segment. The forecasted hours of an identified configuration change will be compared to the actual hours of the same configuration to calculate the difference and the percentage difference. The results will be used to develop a Cost Estimating Relationship (CER) or factor through bivariate regression analysis if feasible.

To answer the second research question, an analysis will determine the number of aircraft produced after a configuration change until the prime contractor was able to return to a stable learning rate. This will be accomplished by removing one production unit at a time (in sequential order beginning with the first unit of the segment) and calculating the learning curve slope of the remaining units until the stable rate of the prior segment is achieved. An overall commonality is not expected because every program, every contractor, and the associated production process are different. Instead, the results are informational and may support contract negotiation efforts with more insight into post-configuration change production.

Conclusion

This chapter explained the methodology used in this study. If the analysis results indicate that a configuration change affects the learning curve rate, a comparison between the impacts may enable the researcher to develop a factor to modify learning curves when a configuration change occurs. This factor could provide DoD cost estimators with another tool to account for the ever changing environment that is DoD acquisition. Chapter IV will walk through the results of the analysis described in this chapter.

IV. Analysis and Results

Introduction

This chapter presents the results of the tests and analysis described in Chapter III. Analysis of the results attempts to answer the research questions previously outlined: is there an impact to the learning curve slope when a configuration change is introduced to the production line? Specifically: what is the learning curve slope for each new configuration; are the production segments for each configuration significantly different; and what is the difference between the hours predicted based on the prior configuration and actual hours for each segment? In addition, how many units of the newly configured aircraft are produced before the contractor regains the stable learning rate? The results of any statistical tests as well as graphical analysis are presented herein.

Normality and Equal Variance

Prior to hypothesis testing, normality and equal variance tests are conducted to determine the appropriate hypothesis testing method between parametric and nonparametric analysis.

The statistical analysis software, JMP 11.0, was used to calculate the Shapiro-Wilk test values in this study. Table 2 summarizes the Shapiro-Wilk test results for the four programs. Small p-values (smaller than 0.05) reject the null hypothesis that the underlying population from which the samples are drawn is normally distributed.

Table 2: Shapiro-Wilk Test of Normality Summary

Program	Test Statistic Value	P-Value
Program A	0.939002	0.0006
Program B	0.929877	0.0686
Program C	0.971010	0.0017
Program D	0.888473	<0.0001

Three of the four programs have a p-value less than 0.05, concluding the data is not normally distributed and parametric analysis is inappropriate. Given that only one of the four programs is normally distributed, nonparametric testing is conducted on all four programs for consistency. Nonparametric testing is more appropriate for this study and none of the tests outlined in Chapter III require equal variance, so meeting that assumption is ignored. The hypothesis also changes to comparing medians as opposed to means:

Hypothesis 1: Is the median amount of labor hours prior to a configuration change the same as the median amount of labor hours subsequent to a production change?

H_0: Median labor hours $_{\text{prior to configuration change}}$ = Median labor hours $_{\text{post configuration change}}$
H_a: Median labor hours $_{\text{prior to configuration change}}$ $\neq$ Median labor hours $_{\text{post configuration change}}$

Visual Analysis

As with any learning curve analysis, this study begins with, and is heavily supported by, visual analysis using scatter-plots. Following the traditional learning curve analysis steps, the analysis began with plotting the data for each program in unit space.

The scatter-plots include only the units identified as when the contractor was substantially into production (as previously defined). The red markers annotate the configuration changes introduced during production on every chart presented. The configuration changes visually demonstrate an impact to the learning curve for each program and break the single pseudo trendline into separate learning curve segments.

The original scatter-plots are then recreated showing the learning curve segments annotated by different colored data points to visually represent the segments that will be used in the remainder of the analysis as well as the individual regression equations for each segmented trendline. The final visual graphs used in the preliminary analysis are scatter-plots of the data in log space. If the data points in log space do not approximate a line, the analysis method should be reconsidered (ICEAA Module 7, 2013:22). Considering different learning curve methods and models is beyond the intent of this research study. Any data segments considered inappropriate under this analysis method based on the log space plots will be excluded from the remainder of the analysis. The following sections provide the visual analysis results for each program.

Program A

Figure 9 shows the initial scatter-plot of the units for Program A. Overall, the data is consistently clustered around the power trendline throughout production. The single red marker indicates the only configuration change identified by the program office for Program A. The contractor incorporated several configuration changes at once, beginning with this single unit. As intuitively expected, the first unit with the new configuration shows an increase in the production labor hours. The scatter-plot shows the contractor resumed learning after the initial introduction of the configuration change and the required touch labor hours decrease as more units are produced. Visually, we can see the learning is at a faster rate (steeper) than the units prior to the configuration change until the units approach the trendline (the contractor's stable learning rate) and the learning flattens again.

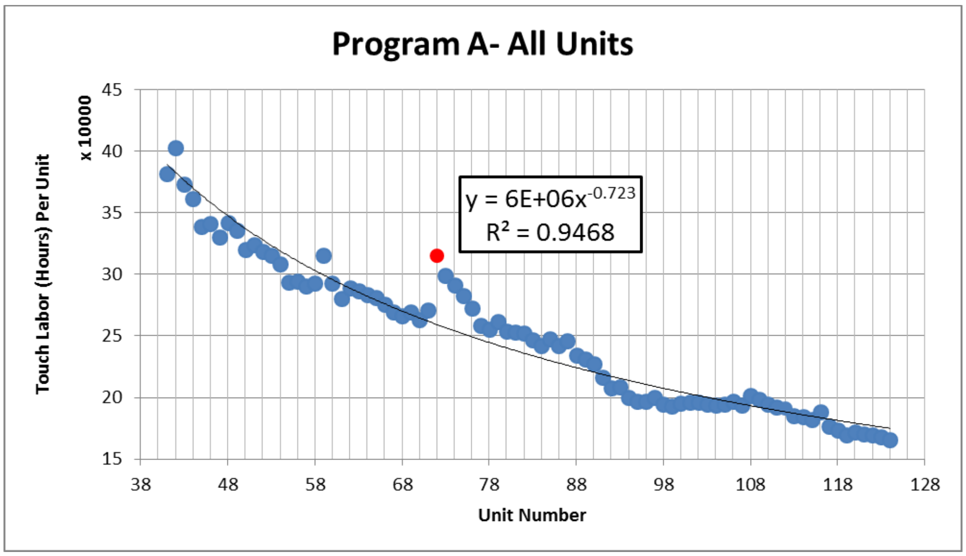

Figure 9: Program A Scatter-Plot (All Units)

Figure 10 shows the recreated scatter-plot for Program A. The units prior to the configuration change are the blue markers (Configuration A) and the units after the change are the maroon markers (Configuration B). The learning curve equations for each segment are shown and individually the R^2 values have decreased for each segment, but both segments still each have trendlines that appear to fit the data very well.

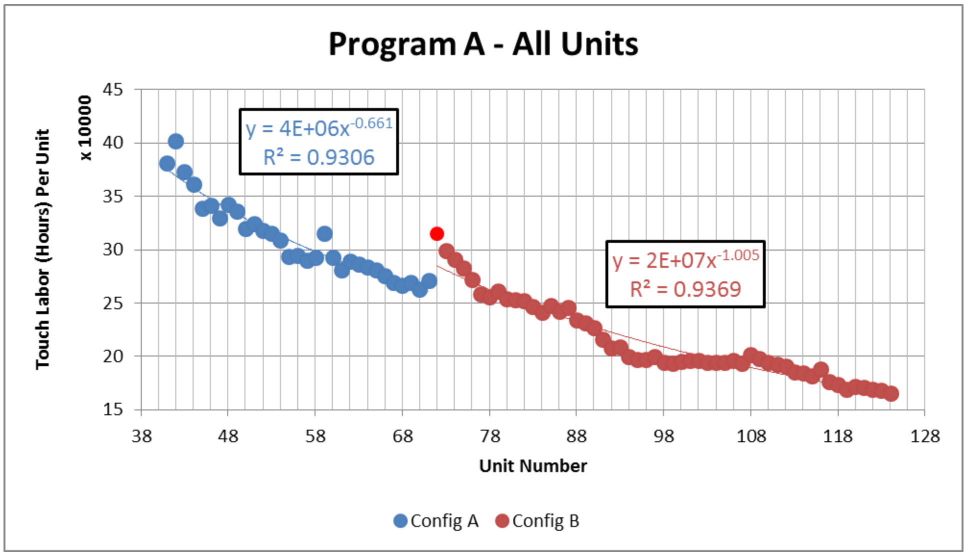

Figure 10: Program A Segmented Scatter-Plot (All Units)

Figure 11 shows the segmented data points for Program A in log space. Both segments appear to approximate a line, so both segments will be included in the next section of analysis.

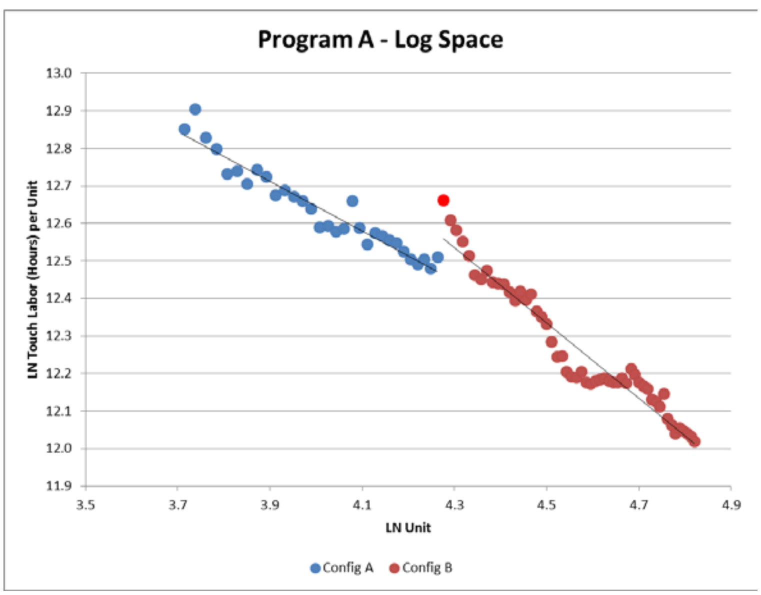

Figure 11: Program A Segmented Scatter Plot in Log Space (All Units)

Program B

Figure 12 shows the initial scatter-plot of the units for Program B. The two red markers represent the two configuration changes identified by the program office for Program B. The single trendline throughout all of the data points does not fit any of the data well, and is representative of the pseudo learning curve phenomenon previously discussed. Without signifying anything on the graph, three distinct segments are visually apparent, indicating a single trendline throughout the data is inappropriate. As intuitively expected, the first unit with the new configuration shows an increase in the production labor hours. The scatter-plot shows the contractor resumed learning after the initial introduction of the configuration change and the required touch labor hours decrease as more units are produced. Counter-intuitively, the first unit of the second configuration change shows a decrease in production labor hours from the previous segment. However, because the total touch labor hours for each unit are less than required for any prior unit in the program, the new configuration may simply be less complicated (requiring less labor) or a more efficient production process may have been created.

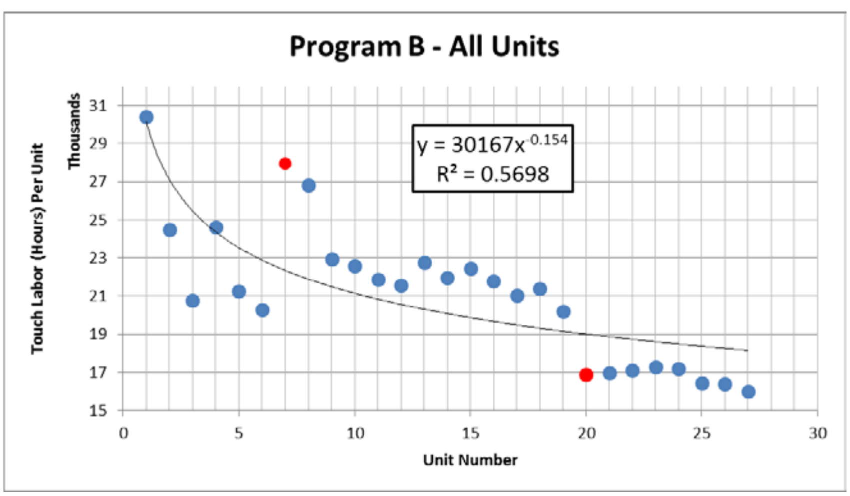

Figure 12: Program B Scatter-Plot (All Units)

Figure 13 shows the recreated scatter-plot for Program B. The units prior to the configuration changes are the blue markers (Configuration A), the units after the first change are the maroon markers (Configuration B), and the units after the second change are the green markers (Configuration C). The learning curve equations for each segment are shown and individually the R^2 values have increased for Configurations A and B and the segmented trendlines appear to fit the data better than the initial single trendline. However, the R^2 value has decreased for Configuration C and the data does not appear to follow a traditional power curve form.

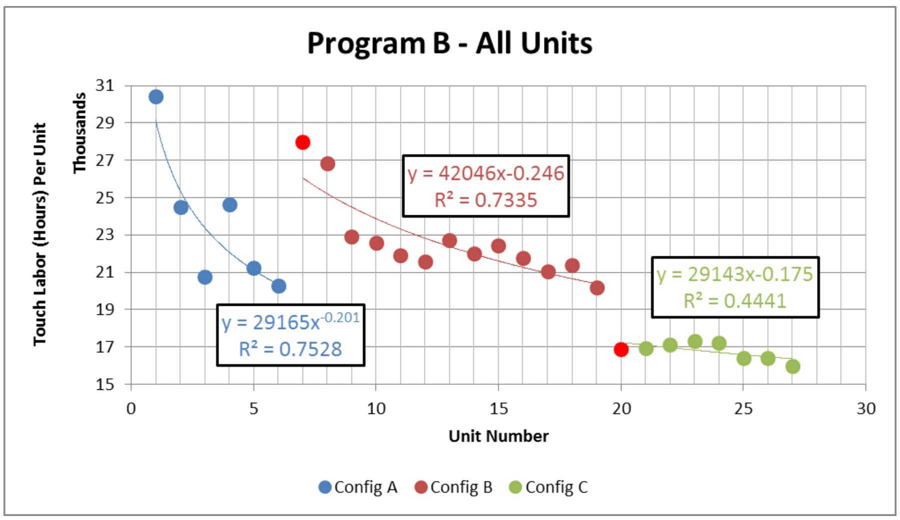

Figure 13: Program B Segmented Scatter Plot (All Units)

Figure 14 shows the segmented data points for Program B in log space. The first and second segments appear to approximate a line, so both segments will be included in the next section of analysis. The third segment clearly does not approximate a line, indicating the analysis method outlined in this study is not appropriate for the third data segment, so the Configuration C segment will be removed for the remainder of the study's analysis.

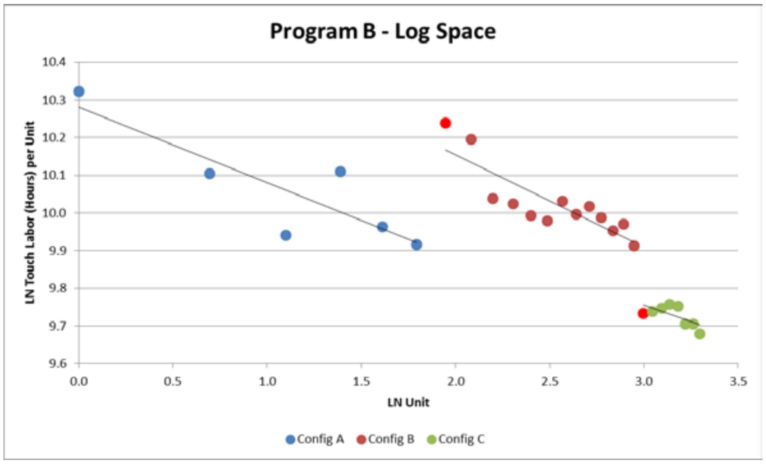

Figure 14: Program B Segmented Scatter Plot in Log Space (All Units)

Program C

Figure 15 shows the initial scatter-plot of the units for Program C. The 15 red markers represent the 15 configuration changes identified by the program office for Program C. The configuration change markers and the impact on the production touch labor hours per unit are sporadic throughout the data points, with no clear discernable patterns due to configuration changes that were seen in the scatter-plots for Programs A and B.

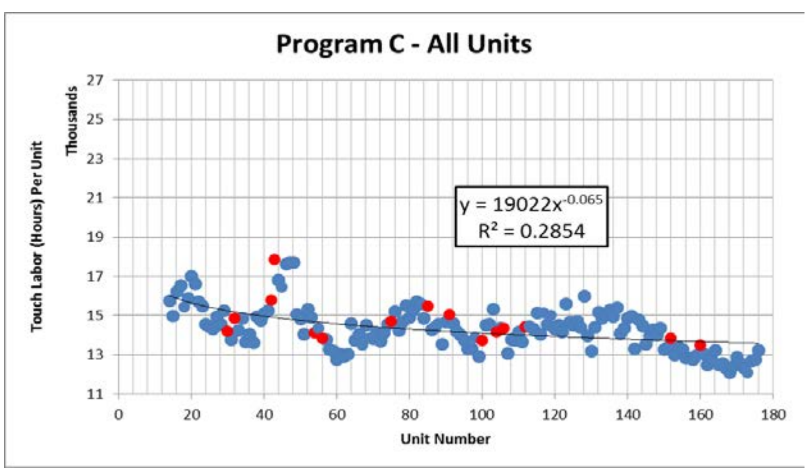

Figure 15: Program C Scatter-Plot (All Units)

Figure 16 shows the re-created scatter-plot for Program C. The researcher attempted to segment the data for Program C to try to identify any apparent trend, which was not discovered. The blue markers indicate four data point segments that best conform to a traditional power curve and visually stand out as possible segments of units that are benefitting from the contractor learning in the production process. The green segments show "un"-learning – overall, the hours required for each subsequent unit are increasing until the next blue segment begins. Because the green segments do not conform to a power curve form, if the log space plot indicates the analysis model of this study is appropriate to model the data points, additional steps will be required to analyze this data set.

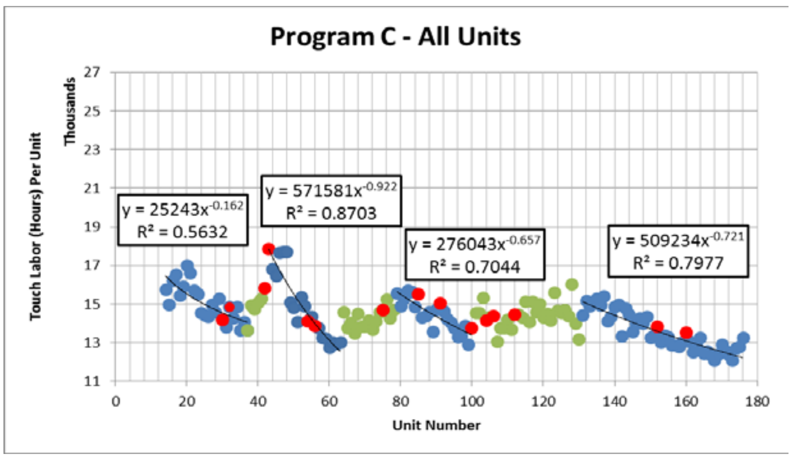

Figure 16: Program C Segmented Scatter-Plot (All Units)

Figure 17 shows the data points for Program C in log space. None of the data appears to approximate a line, indicating the analysis method outlined in this study is not appropriate, so Program C will be removed for the remainder of the study's analysis. Possible reasoning for the sporadic behavior of Program C's data will be explored in Chapter V.

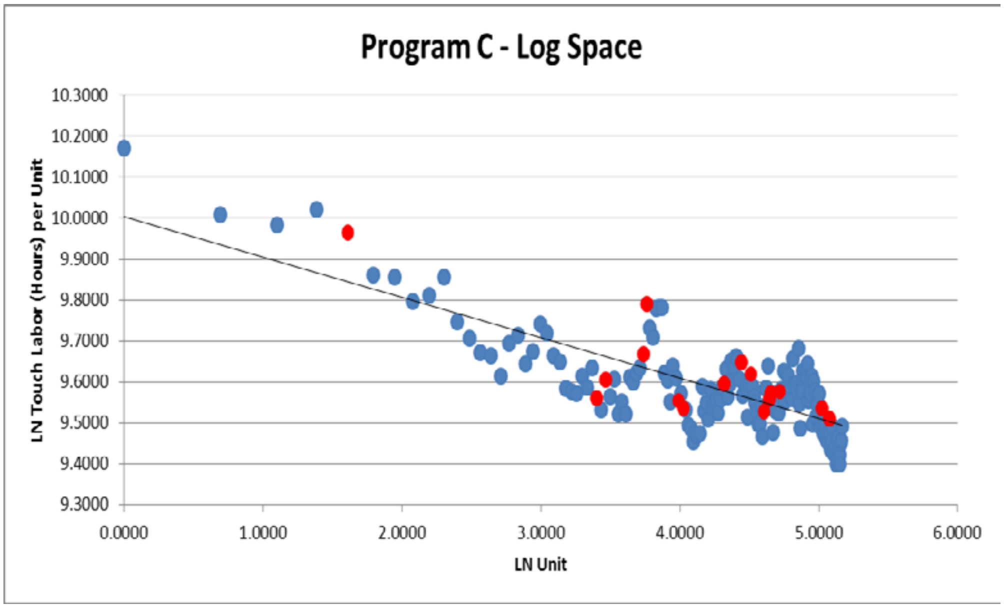

Figure 17: Program C Scatter-Plot in Log Space (All Units)

Program D – Final Assembly

Figure 18 shows the initial scatter-plot of the units for Program D. The initial data of Program D is split between three variants because this program is produced in three different modifications. Only the final assembly is conducted on the same production line, so only the final assembly learning can be analyzed to meet the intent of this research study even though the data trend at the final assembly level is very similar to the data trend of the total touch labor hours per unit. These variants are not the configurations in which the data will be analyzed.

The five red markers indicate the five configuration changes identified by the researcher to segment the data. Variant B always requires more touch labor hours than Variant A and Variant C always requires more touch labor hours than Variant B. The segments were chosen based on when Variants B and C were introduced into the final assembly production line. Towards the end of production, Variants B and C together were overwhelmingly separated from Variant A, so two separate segments were identified to best fit the data and not establish a pseudo learning curve through the center of all the data points.

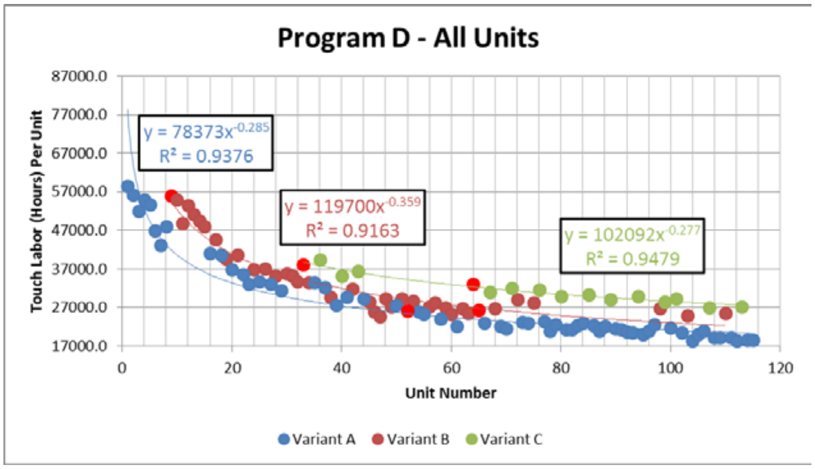

Figure 18: Program D Scatter-Plot (All Units)

Figure 19 shows the re-created scatter-plot for Program D and better illustrates the naturally occurring segments present in the data. The five points identified as configuration changes form six data segments, signified in the figure.

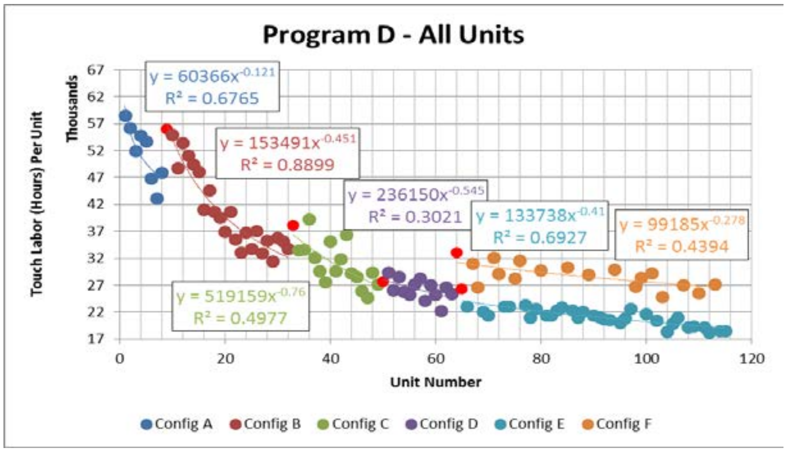

Figure 19: Program D Segmented Scatter-Plot (All Units)

Figure 20 shows the segmented data points for Program D in log space. Each segment appears to approximate a line and will remain in the analysis.

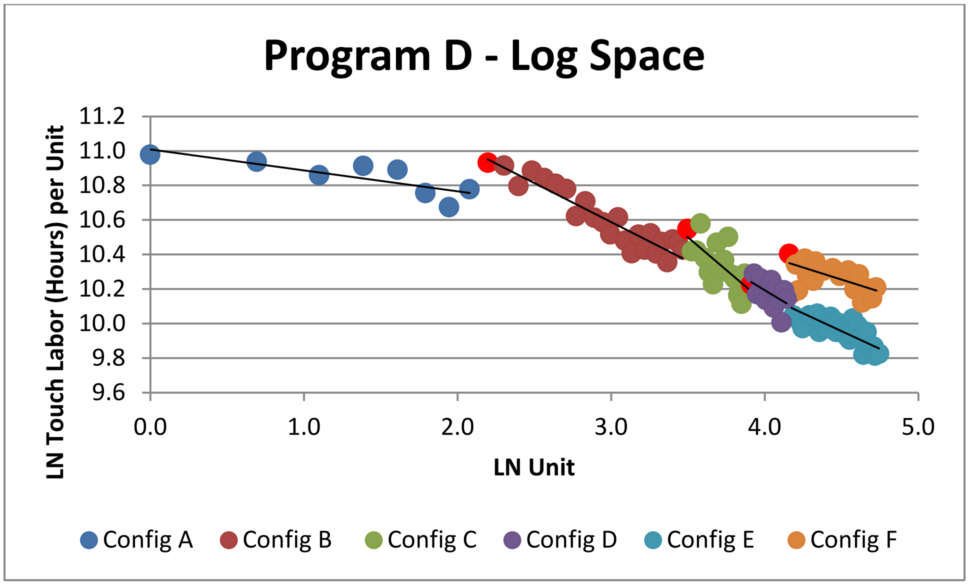

Figure 20: Program D Segmented Scatter-Plot in Log Space (All Units)

Statistical Analysis

The next phase of analysis involves the nonparametric statistical analysis to determine if the segments remaining in the analysis are statistically similar. Table 3 includes the slope calculations for each program for each segment identified. Configuration A is always the initial configuration, prior to any changes. Based on this summary, the slope never remained the same after a configuration change.

Table 3: Segment Learning Curve Slope Values

		Configuration					
		A	B	C	D	E	F
Program	A	63.26%	49.84%	-	-	-	-
	B	87.02%	84.33%	-	-	-	-
	D	91.96%	73.15%	59.04%	68.56%	75.24%	82.49%

The configuration changes appear to have an impact to the slope, but statistical testing is required to understand if the change in slope may be a contributing factor causing the segments to be statistically different. The Wilcoxon Rank Sum Test was conducted on Program A because it is most appropriate to compare two samples when both sample sizes are large. The Mann-Whitney U Test was conducted on Program B because it is most appropriate to compare two samples when the sample sizes are not large (both less than 30). The Kruskal Wallis Test was first conducted on Program D because it is most appropriate to compare three or more samples.

The Kruskal Wallis Test for Program D calculated a K value of 98.938 and a critical value of 11.070. Because the critical value is less than the K value, the null hypothesis that the medians for all of the segments are statistically similar is rejected. The Kruskal Wallis test results reveal that at least one of the segments is statistically different, but does not indicate which segment(s). A combination of the Wilcoxon Rank Sum Test and Mann-Whitney U Test was also conducted on successive pairs of segments to determine which segment(s) were statistically different from the others for Program D. Successive comparison is the most appropriate because the intent of this study is to determine if a single configuration change influences the following production units. If there were another configuration change introduced in subsequent learning curve segments, comparing the first to the last would be inappropriate because more than one configuration change is affecting the unit touch labor hours, and the impact from one to the next cannot be isolated. The Mann-Whitney test was used when the configuration pairs both had small samples (less than 30) and the Wilcoxon test was used when at least one sample size was greater than 30.

Program A Results

The Wilcoxon Rank sum test was performed on Program A because Configuration A has a sample size of 32 and Configuration B has a sample size of 53 (both large). With large sample sizes, z-values can be used to analyze the results. At a 95% confidence level, the corresponding z-score is 1.96. The test statistic z-value for Program A was 7.157. Because the absolute value of the test statistic is greater than 1.96, the null hypothesis is rejected, signifying the medians between the two samples are statistically different.

Program B Results

The original data for Program B included three configurations. The Kruskal Wallis test was performed to determine if at least one of the sample medians was statistically different from the others. The Kruskal Wallis Test for Program B calculated a K value of 16.286 and a critical value of 5.991. The null hypothesis that the medians for all of the segments are statistically similar is rejected because the critical value is less than the K value.

Through prior analysis, the third configuration, Configuration C has been excluded, but the Mann-Whitney U Test was still performed on both pairs of segments to identify which segment median(s) are statistically different. For the first pair, Configurations A and B, the calculated U statistic value is 39. The table critical value for these sample sizes (six and thirteen, respectively) is 16. The results fail to reject the null hypothesis because the U statistic is greater than the critical value, signifying the median values for these configuration segments are statistically similar.

For the second pair, Configurations B and C, the calculated U statistic value is zero. The table critical value for these sample sizes (thirteen and eight, respectively) is 24. Because the U statistic is less than the critical value, the null hypothesis is rejected; signifying the median values for these configuration segments is statistically different. Due to the reasons stated previously in the analysis, Configuration C, although statistically different from Configuration B, is not appropriately modeled through the analysis in this study.

Program D Results

Table 4 summarizes the sample sizes for each segment in Program D.

Table 4: Program D Segment Sample Size Summary

Segment	Sample size (n)
Configuration A	8
Configuration B	24
Configuration C	17
Configuration D	14
Configuration E	33
Configuration F	19

The Mann-Whitney test was performed on the first three segment pairs (between Configurations A, B, and C) because all there sample sizes are less than 30. The Wilcoxon test was performed on the last two segment pairs (between Configurations D, E, and F) because Configuration E has a sample size of greater than 30 and is the configuration compared with both Configurations D and F. Table 5 summarizes the test results.

Table 5: Program D Mann-Whitney and Wilcoxon Test Results Summary

Segment Pairs	Mann-Whitney U-Statistic	Mann-Whitney Critical Value	Wilcoxon Test Statistic z-value	95% Confidence Level comparison z-score	Decision
A and B	29	50	-	-	Reject H_0
B and C	46	129	-	-	Reject H_0
C and D	31	67	-	-	Reject H_0
D and E	-	-	5.03	1.96	Reject H_0
E and F	-	-	5.81	1.96	Reject H_0
Mann-Whitney Analysis: Reject H_0 when U-statistic < Critical Value Wilcoxon Analysis: Reject H_0 when Test Statistic z-value > Confidence level z-score					

The results in this table indicate that every segment is statistically different from the adjacent segment for Program D.

Investigative Questions Answered

The previous section provided evidence that in nearly every case involved in this study, the segmented data are statistically different when comparing adjacent segments, which addresses that issue in the first research question. There is a change to the learning curve slope each time a configuration change is introduced, and in every case analyzed except one, the median labor hours (which are partially a function of the learning curve slope) for the different configurations is statistically different. These findings statistically support that using the prior learning curve equation is inappropriate to predict the hours of the new configuration because the units come from different populations.

Further addressing the first set of questions, the learning curve equation for each segment is used to predict the touch labor hours for each unit in the following segment. The total predicted hours for each segment are compared to the total actual hours of the segment and the results are shown as a difference in hours as well as a percent difference for comparison between the programs. A negative value indicates the estimate was lower than the actuals. Table 6 details the results of the predicted and actual hour comparisons.

Table 6: Learning Curve Equation Prediction vs. Actuals Summary

Program A				
	Predicted Hours	Actual Hours	Difference	% Difference
A predicting B	11,336,756.40	11,371,252.00	(34,495.60)	-0.30%
Program B				
	Predicted Hours	Actual Hours	Difference	% Difference
A predicting B*	229,114.62	295,348.35	(66,233.73)	-22.43%
Program D				
	Predicted Hours	Actual Hours	Difference	% Difference
A predicting B	1,014,525.48	986,331.30	28,194.18	2.86%
B predicting C	490,909.41	531,988.54	(41,079.13)	-7.72%
C predicting D	339,726.00	368,921.32	(29,195.31)	-7.91%
D predicting E	678,070.58	698,789.63	(20,719.06)	-2.96%
D predicting F	397,530.17	542,429.97	(144,899.80)	-26.71%
*Configuration B not considered a statistically significant change from configuration A				

Given that this portion of the analysis only includes three programs, and two of the programs only compare two segments, there are too few data points to develop any meaningful CER or factor. However, the results are still impactful because for each of the seven segment comparisons, no fewer than 20 thousand hours was the difference between the predicted and actuals, which equates to millions of dollars per segment misestimated (generally underestimated) in a cost estimate. Underestimation requires the program office to find dollars not currently in its budget and overestimation temporarily ties up funding that can be used for other purposes.

In reality, a contractor will submit a tech-refresh proposal to the program office to account for the configuration change, but will estimate the unit costs based on an extrapolation of its stable learning curve because the new slope is unknown. In every program analyzed in this study, the learning curve slope becomes much steeper after the configuration change (when compared to the initial stable slope), and a extrapolation of the stable curve will create a higher per unit cost than the contractor would actually experience with the steeper learning curve. This phenomenon is explored in the next section, which analyzes Program A to answer the second research question of how many production units are manufactured before the contractor returns to its stable learning rate.

Program A was selected for analysis in answering the second research question because Program A has a large sample size in total and in each segment separately. In addition, only one configuration change came into the production line, so this program provides the simplest situation to analyze. The stable slope for Program A is 63.26% as determined by the units in Configuration A (units 41 to 71). Table 7 summarizes the slopes for Configuration B beginning with units 72 to 124 and removing one unit at a time from the beginning of the segment until the stable slope was reestablished.

Table 7: Program A Stable Slope Analysis Summary

First Unit	Slope	Units to Stabilize
72	49.84%	
73	50.69%	1
74	51.34%	2
75	51.95%	3
76	52.48%	4
77	52.85%	5
78	52.83%	6
79	52.81%	7
80	53.21%	8
81	53.44%	9
82	53.80%	10
83	54.36%	11
84	54.85%	12
85	55.25%	13
86	56.33%	14
87	57.39%	15
88	59.18%	16
89	60.52%	17
90	62.03%	18
91	63.60%	19
92	64.36%	
93	64.30%	
94	64.52%	
95	63.51%	
96	62.06%	
97	60.54%	

The stable learning rate is achieved with the production of unit 91, which are 19 units after the configuration change came into the production line. While every program will stabilize at different production rates, the important point in this analysis is that after the configuration change is introduced, the contractor learns much quicker on the units after the configuration change than the stable flatter learning rate pre-change. The units immediately following the stabilized rate (92 to 97) are included in the table to show that the contractor does not continue to learn for all units after the stabilized rate is achieved, rather the contractor's learning rate stays around the stabilized rate. While this analysis is for only one program and cannot be generalized for all programs, the prior analysis did show that for each program, the contractor learned at a much steeper rate following the configuration change. These results do provide evidence to support a position other than the contractor extrapolating the prior stable learning curve in a tech-refresh proposal before a configuration change is introduced.

Split Learning Curves

Chapter II introduced the idea of split learning curves and a possible formula to add the new learning curve slope of the units post configuration change to the learning curve slope of the prior segment to create a single learning curve equation. The formula was caveated with the statement that the equation is best demonstrated in a situation with low production, like satellites or ship building. The data available for this study was not in a form to investigate the usefulness of the equation in a situation with greater amounts of production like aircraft manufacturing. The delta in hours attributed to the configuration change itself is required for the equation, which was not available during this analysis.

Summary

The purpose of this chapter was to provide the analytical results based on the methodology described in Chapter III. While there were not enough programs with data to include in this analysis to generalize results or create an adjustment factor, the results show there is possibly a significant impact to the learning curve slope. The slope changing for each configuration, the statistical difference in the segments, which are partially a function of the slope, and the difference in estimated and actual hours may support that theory. Chapter V will summarize the significance of this research study as well as recommendations based on this research and recommendations for future research.

V. Conclusions and Recommendations

Introduction

This research study started with two main research questions. First, to determine if a configuration change brought into an aircraft production line had a significant impact on the production learning curve slope. Specifically: what is the learning curve slope for each new configuration; are the production segments for each configuration significantly different; what is the difference between the predicted and actual hours for each segment? Second, to determine after a configuration change is introduced, how many production units are manufactured before the contractor's stable learning rate is again achieved. Hypothesis testing was used to assess the statistical differences in the median labor hour values pre- and post-configuration change to determine the presence of a statistically significant impact possibly due to an unstable learning curve. Data point analysis was used to address the second research question. The raw results from the hypothesis testing and analysis are shown in Chapter IV. This Chapter V addresses the impacts of the findings and any conclusions that can be drawn as well as the research limitations, significance, and recommendations for future research.

Conclusions of Research

The hypothesis testing indicated a statistically significant difference in the median production touch labor hours in the pre-configuration change and post-configuration change aircraft for every pair of data segments analyzed, except for one. Comparing the median values may equate to a statistically significant difference in the learning curve slopes for those data segments because the impact to the learning curve slope is evaluated through the touch labor hours of the data points, as they are partially a function of the slope value.

The data point analysis to address the stable learning curve research question produced interesting results. The analysis did show a pattern that post-configuration change, the contractor initially learns at a much faster rate and the learning rate decreases with each subsequent unit until the stable learning rate is again achieved. The learning rate did appear to stabilize at this point and did not continue to decrease.

While data for too few programs was available at the time of this research study to generalize the results or develop an adjustment factor, the results of this study may imply two things. First, that a majority of the time there is an impact to the learning curve slope whenever a configuration change is introduced during production. Second, that the contractor is able to learn to incorporate the change much more quickly than its stable learning rate for the entire aircraft. However, unless more data is studied, these results cannot be generalized in any way.

Significance of Research

The results of this research indicate there may be a significant impact to the learning curve slope when a configuration change is introduced during production, even if the program is substantially into production, as were the programs included in this analysis. While the results cannot be generalized, the findings suggest more research in this area is important for two reasons. First, if more programs are analyzed, more data points may lead to the development of a CER or factor to adjust a stable learning curve, which would be a useful tool for cost estimators given the ever changing acquisition environment. Second, because the learning curve slope is such a crucial factor in production contract negotiations, empirical evidence strengthens the DoD's position of what the contractor's expected learning curve should be – which this study has not found to be an extrapolation of the contractor's stable learning curve.

An initial estimate that does not anticipate any configuration changes will underestimate unit production hours or costs required for the newly configured unit. If the DoD negotiates a contract based on an extrapolation of the contractor's stable rate, these results provide evidence that the stable rate will over estimate the production requirements; this analysis showed the contractor learns at a steeper rate after a configuration change. The initial underestimating, coupled with the contractor's overestimation, will result in the program office requesting millions of dollars, possibly in excess, per configuration change.

Learning curve theory advises the use of the most recent or most representative production articles to predict the follow on articles. While this is intuitive and proven to result in better estimates even in this study, program offices cannot disregard the prior units. If program offices track the configuration change information and the resulting impacts, the DoD may be in a better position to estimate costs and negotiate production contracts.

Assumptions and Limitations

The overarching methodology assumptions may lead to different results if a different methodology was chosen. This study was designed to evaluate the impact to the learning curve only when a configuration change was identified. If the data points were grouped or segmented based on a different methodology, different results may be uncovered. The methodology also depends on the program office accurately and fully identifying configuration changes among the data set.

The overarching limitation in this research study was the availability of data at the necessary level to meet the intent of the study. With so few programs available to analyze at the time this study was conducted, no generalizations can be made regardless of any statistical significance. A limitation to the analysis was not having insight into the contractor's production process changes that may also attribute to the change in production hours. Intuitively, a configuration change could result in a production process change at the same time, which may affect the data point values. Two subject matter experts believe technology cycle and production process changes could be the reason the configuration changes did not appear to cause the only change to the learning for Program C.

Recommendations for Future Research

There are five potential areas for future research that should be considered. First, conducting the same analysis on more programs would provide additional empirical evidence of what happens to a learning curve when a configuration change is introduced into production, which the benefits of have already been explained. Second, an extension of conducting the analysis on more programs is to analyze more aircraft class types (multiple of each type when possible) and to obtain data about the size and type of the configuration change. Figure 21 depicts a possible research model with moderating effects.

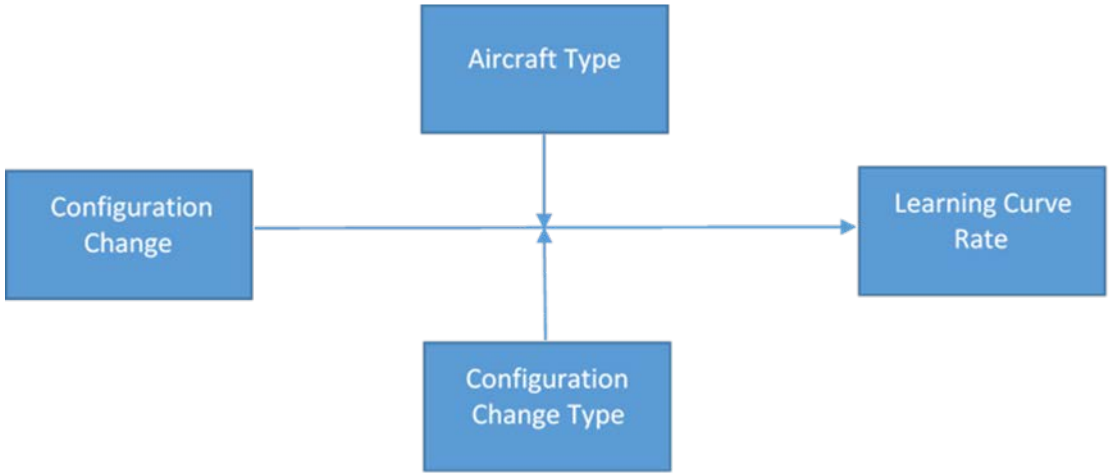

Figure 21: Possible Future Research Model with Moderating Effects

The inclusion of these moderating effects (and other identifiable effects) may allow the development of a statistical model for program office use. Third, including an incorporation of any learning curve models and elements discussed in Chapter II may better analyze the impact of different factors to the learning curve simultaneously. Fourth, investigation into contractor production process changes in preparation for or conjunction with configuration changes (Program C) could develop a way to model these situations. Finally, studying the impact to the learning curve slope of missiles rather than aircraft is a possible future area of research worth exploring. The DoD procures many more missiles than aircraft, so the sample sizes will be much larger if the data is attainable, which may lead to more conclusive and generalizable results.

Bibliography

Anderlohr, G. “Determining the Cost of Production Breaks.” *Management Review*, 16-19 (September 1969)

Avinger, C. R. *Analysis of Learning Curve Fitting Techniques*. MS Thesis, AFIT/GSM/LSQ/87S-3, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1987.

Badiru, A. B. “Computational Survey of Univariate and Multivariate Learning Curve Models.” *IEEE Transactions on Engineering Management*, 39(2): 176-188. (1992).

Badiru, A. B. “Multivariate Analysis of the Effect of Learning and Forgetting on Product Quality.” *International Journal of Production Research*, 33(3): 777-794 (1995).

Bailey, C.D. “Forgetting and the Learning Curve: A Laboratory Study.” *Management Science*, 35(3): 340-352 (1989).

Daulhaus, F. J., and Roj, J.S. Department of the Army. *Learning Curve Methodology for Cost Analysts*. Washington DC. U.S. Department of the Army. 1967.

Department of the Air Force. *Air Force Cost Estimating Handbook.* Washington D.C. U.S. Department of the Air Force. 2007.

Ford, C. (n.d.). “A Rule of Thumb for Unequal Variances.” *Data.library.virginia.edu* Retrieved on January 13, 2015.

International Cost Estimating and Analysis Association. *Cost Estimating Body of Knowledge Training, Unit I, Module 1. Cost Estimating Basics*. 2013.

International Cost Estimating and Analysis Association. *Cost Estimating Body of Knowledge Training, Unit III, Module 7. Learning Curve Analysis: How to Account for Cost Improvement*. 2013.

Kronemer Alexander and Henneberger, J. Edwin. “Productivity in Aircraft Manufacturing.” *Monthly Labor Review*, 24-33 (June 1993).

Lam, K.C., Lee, D., and Hu, T. “Understanding the Effect of the Learning-Forgetting Phenomenon to Duration of Project Construction.” *International Journal of Project Management,* 19(7): 411-420 (2001).

McClave, J. T., Benson, P. G., and Sincich, T. *Statistics for Business and Economics (12th ed.).* Boston, MA: Pearson Education, Inc., 2014.

Searle, David A. *The Impact of the Packard Commission's Recommendations on Reducing Cost Overruns in Major Defense Acquisition Programs*. MS thesis, AFIT/GCM/LAS/97S-12. School of Logistics and Acquisition Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1997.

Tracht, Tom. *An Analysis of the Impact of Log-Linear Regression on the Estimated Learning Curve Parameters*. MS thesis, AFIT/GCA/LSQ/88S-9. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1988.

Yelle, Louis E. "The Learning Curve: Historical Review and Comprehensive Survey." *Decision Sciences*, 10: 302-328 (1979).

Watkins, P. N. "The Persistence of Learning and Acquisition Strategies." *Acquisition Review Quarterly*, 15-30 (2001).

Wright, T.P. "Factors Affecting the Cost of Airplanes." *Journal of Aeronautical Sciences*, 3: 122-128 (February 1936).

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704–0188*
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704–0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.			
1. REPORT DATE (DD–MM–YYYY) 24-03-2016	2. REPORT TYPE Master's Thesis		3. DATES COVERED (*From — To*) Sept 2014 – March 2016
4. TITLE AND SUBTITLE An Analysis of the Impact of Configuration Changes to the Learning Curve for Department Of Defense Aircraft Acquisition Programs Substantially Into Production			5a. CONTRACT NUMBER
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Honious, Candice M, Civ			5d. PROJECT NUMBER
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of 2950 Hobson Way WPAFB OH 45433-7765			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENV-MS-16-M-256
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Cost Analysis Agency Mr. Anthony Munafo 2275 D Street WPAFB OH 45433 Anthony.Munafo@us.af.mil			10. SPONSOR/MONITOR'S ACRONYM(S) AFCAA
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION / AVAILABILITY STATEMENT Distribution Statement A. Approved for Public Release; Distribution Unlimited			
13. SUPPLEMENTARY NOTES This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States			
14. ABSTRACT This research is one of the first empirical studies of the impact to the learning curve slope and touch labor hours when production is continuous but a configuration change interrupts the learning process. The analysis discovered the learning curve slope after a configuration change is different from the stable learning curve slope pre-configuration change. The differently configured units were found to be statistically different from one another, which may be due to the unstable slope, given that the labor hours per unit are partially a function of the learning rate. The significant difference between the configurations provides statistical support that the new configuration should not be estimated with the learning curve equation of the prior configuration. The research also discovered the post-configuration slope is always steeper than the stable learning slope. Therefore, estimating the new configuration based on the slope of the units pre-configuration change will result in over-estimation, but an initial estimate with a stable slope and no anticipated changes will under-estimate the production hours once a change is required. The steeper slope decreased with each subsequent unit until the slope stabilized.			
15. SUBJECT TERMS Learning curve, configuration change, slope impact, aircraft production			

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Dr. John Elshaw, AFIT/ENV
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	72	19b. TELEPHONE NUMBER (Include Area Code) (937) 255-3636 x4650 John.Elshaw@afit.edu